JOURNAL OF VIROLOGY, Oct. 1998, p. 7985–7991 Vol. 72, No. 10
0022-538X/98/$04.00+0

A Virus-Encoded RNA Polymerase Purified from Baculovirus-Infected Cells

LINDA A. GUARINO,[1,2,3]* BIN XU,[1] JIANPING JIN,[1] AND WEN DONG[2]

Departments of Biochemistry and Biophysics[1] and Entomology[2] and Center of Advanced Invertebrate Molecular Sciences,[3] Texas A&M University, College Station, Texas 77843-2128

Received 16 March 1998/Accepted 25 June 1998

A DNA-dependent RNA polymerase was purified to homogeneity, starting from insect cells infected with the baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV). The purified polymerase supported accurate and specific transcription from late and very late promoters but was not active on viral early promoters. Thus, promoter recognition is an integral function of the purified enzyme. The purified RNA polymerase was composed of only four equimolar subunits, which makes it the simplest DNA-directed RNA polymerase from a eukaryotic source described so far. Amino-terminal protein sequencing, peptide fingerprinting, and immunochemical analyses were used to identify the four subunits, all of which are virus encoded. Overexpression of the four viral proteins (LEF-8, LEF-4, LEF-9, and p47) in baculovirus-infected cells resulted in a significant increase in the levels of RNA polymerase produced in the infected cells. Thus, the overexpression data are consistent with our identification of the RNA polymerase subunits.

Autographa californica nuclear polyhedrosis virus (AcNPV) is the prototype member of the *Baculoviridae*, a large family of DNA viruses that are pathogenic for invertebrates. AcNPV has a circular DNA genome of 134 kb and potentially encodes 154 proteins (1). Like many DNA viruses, AcNPV gene expression is temporally regulated. Both immediate-early and delayed-early genes are expressed before viral DNA replication, while late and very late gene expression is dependent on viral DNA replication (4). Unlike other eukaryotic DNA viruses that replicate in host nuclei, baculoviruses use two different RNA polymerases for transcription of their genes. The early viral genes are transcribed by host RNA polymerase II (12, 18), and the early promoters contain motifs common to RNA polymerase II promoters (16, 24). The late and very late genes, however, are transcribed by an RNA polymerase that is resistant to α-amanitin and is chromatographically distinct from the three host RNA polymerases (15, 31). In addition, the baculovirus late promoters lack motifs characteristic of promoters recognized by all three host polymerases. The nucleotide sequences that are essential for late transcription have been mapped to a conserved 12-bp motif surrounding the start of transcription (16, 25, 26). Taken together, these data suggest that the virus encodes its own RNA polymerase. However, it has proven difficult to purify the virus-induced polymerase to homogeneity (3, 31). Surprising, the complete sequence of the viral genome failed to reveal proteins with extensive homology to other DNA-dependent RNA polymerases (1), although limited sequence similarities were noted for two viral proteins, LEF-8 and LEF-9 (19, 23).

To better understand the mechanisms of transcriptional regulation of the AcNPV late and very late genes, two approaches have been used. Todd et al. (29) used a transient expression assay to identify 18 viral proteins involved in the temporal expression of late and very late genes. Approximately half of these are required for early gene expression and DNA replication, while the remaining are potential candidates for the virus-encoded RNA polymerase and associated transcription factors. We and others (13, 20, 30) have developed in vitro transcription systems for the AcNPV late promoters. We constructed transcription templates containing cytidine-free cassettes linked to either the *39k* late promoter or the polyhedrin (*polh*) very late promoter (30). Here we report the use of this template-specific assay to purify the RNA polymerase responsible for transcription of viral late genes. The RNA polymerase is a complex of four virus-encoded proteins and has both promoter recognition and catalytic activities.

* Corresponding author. Mailing address: Department of Biochemistry and Biophysics, Room 103A, Texas A&M University, College Station, TX 77843-2128. Phone: 409-845-7556. Fax: 409-845-9274. E-mail: lguarino@bioch.tamu.edu.

MATERIALS AND METHODS

Preparation of nuclear extracts. *Spodoptera frugiperda* (Sf9) cells were cultured and infected with the E2 strain of AcNPV as previously described (28). Nuclear extracts were prepared from AcNPV-infected Sf9 cells at 36 h postinfection, with two modifications to the previous protocol (30). After Dounce homogenization, the nuclei were washed twice by low-speed centrifugation in hypotonic buffer containing 6% sucrose and then pelleted through a 30% sucrose cushion by centrifugation at 3,000 × *g* for 10 min. After centrifugation of the nuclear extracts at 100,000 × *g*, the supernatants were frozen in liquid nitrogen and stored at −80°C.

Purification of the RNA polymerase complex. All procedures were carried out at 4°C. Nuclear extracts were prepared from two to five 1-liter cultures and frozen until a total of 25 liters of infected cells had been collected. Pooled nuclear extracts were treated with 0.1% polymin P to precipitate nucleic acids. Soluble proteins were then precipitated with 50% ammonium sulfate. The ammonium sulfate precipitate was collected by centrifugation, resuspended in 25 ml of buffer A (50 mM Tris [pH 7.9], 0.1 mM EDTA, 1 mM dithiothreitol) containing 0.5 M $(NH_4)_2SO_4$, and loaded onto a 12-ml phenyl-Sepharose column (Pharmacia) at a rate of 2 ml/min. The column was washed with 50 ml of buffer A–0.5 M $(NH_4)_2SO_4$, and bound protein was eluted with buffer A–300 mM KCl–0.5% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS; Pierce). Fractions (1.8 ml) were collected and assayed for transcription activity. Active fractions were pooled, dialyzed against buffer A–300 mM KCl–0.1% CHAPS, and applied at 1 ml/min to a 5-ml heparin (Bio-Rad) column connected to a Pharmacia FPLC system previously equilibrated with buffer A–300 mM KCl–0.1% CHAPS. The column was washed with loading buffer and eluted with a 20-ml linear gradient from 300 to 500 mM KCl. Peak fractions were pooled, the KCl concentration was adjusted to 250 mM, and the protein was applied to a Mono Q HR 5/5 column (Pharmacia) previously equilibrated with buffer A–200 mM KCl–0.1% CHAPS. The column was washed with 10 ml of loading buffer and then eluted with a 20-ml linear KCl gradient from 200 to 500 mM. Fractions that contained transcription activity were concentrated to 200 μl and filtered through a Superose 6 column in buffer A–2 M KCl–0.1% CHAPS. Fractions (0.5 ml) were individually dialyzed against buffer A–250 mM KCl–0.1% CHAPS, assayed for transcription activity, frozen in liquid nitrogen, and stored at −80°C. Protein concentrations of the crude extract and partially purified fractions were determined by the method of Bradford (6). The concentration of the purified

complex was determined by UV absorbance using a molar extinction coefficient of 259,350, predicted by the amino acid sequences of the four subunits (1).

In vitro transcription assays. In vitro transcription assays for AcNPV late and very late promoters were performed by using slight modifications of conditions previously described (30). Transcription reaction mixtures contained 50 μg of nuclear extract or 10 μl of column fractions and the following components in a volume of 50 μl: 25 mM Tris (pH 7.9), 100 mM KCl, 2 mM $MgCl_2$, 1 mM dithiothreitol, 1.0 mM each ATP and UTP, 20 μM GTP, 5 μCi of [α-^{32}P]GTP (800 Ci/mmol), 5 U of RNasin, 0.2 U of inorganic pyrophosphatase, and 1.0 μg each of Polh/CFS and 39kL/CFS (described in Results). Addition of α-amanitin was not necessary and was not routinely used. As shown by Xu et al. (30), the host RNA polymerases cannot transcribe the baculovirus templates used in these assays. Components were added to the enzyme at the same time, reaction mixtures were incubated for 12 min at 30°C, and then the reaction was stopped by the addition of 150 μl of stop buffer (50 mM Tris [pH 7.5], 1% sodium dodecyl sulfate [SDS], 5 mM EDTA, 25 μg of tRNA per ml). RNA was extracted once with phenol-chloroform (1:1), precipitated with ethanol, resuspended in 90% formamide, and resolved on a 6% polyacrylamide–8 M urea gel. For quantitation of RNA polymerase activity, transcription reaction mixtures were spotted onto glass fiber filters and precipitated with trichloroacetic acid. By definition, 1 U of transcription activity incorporates 1 pmol of GMP into RNA in 30 min at 30°C. The conditions for in vitro transcription from early viral promoters have been previously described (32).

Peptide sequencing. Purified RNA polymerase was concentrated, loaded into two wells of an SDS–8% polyacrylamide gel, and then electrophoresed. Proteins were transferred to polyvinylidene difluoride membranes, stained with Ponceau S, and submitted to the Protein/Peptide Micro Analytical Laboratory (California Institute of Technology). Samples were submitted to automated Edman degradation on an Applied Biosystems 476A sequencer. N-terminal amino acid sequences were obtained for the two smallest proteins. No useful sequence information was obtained for the two larger proteins.

Trypsin digestion followed by mass analysis was used for identification of the second-largest protein. After electrophoresis on an SDS-polyacrylamide gel (SDS-PAGE), proteins were stained in 0.1% Coomassie blue in water. The protein was excised, and subsequent digestion with trypsin and mass analysis were conducted at the Protein/Peptide Micro Analytical Laboratory (California Institute of Technology). Mass spectrometry was performed on a Perseptive Biosystems Elite matrix-assisted laser desorption ionization time-of-flight (MALDI TOF) apparatus. Tryptic peptides masses were used by the MOWSE program to search a peptide mass database constructed from a theoretical trypsin digest of all proteins in the OWL database (21). Search parameters used were a molecular weight filter of 25%, a mass tolerance of 0.1%, and a partial cleavage score factor of 0.2.

Preparation of LEF-8 antiserum. The AcNPV *lef-8* gene was amplified by PCR using an upstream primer (5′-AATCGCTTCCATATGACGGACGTGGTTCAAG-3′) which produced an *Nde*I site (underlined) at the translation initiation codon of the *lef-8* coding sequence and a second primer (5′-GTTTGCAATCGTGCAAGC-3′) that hybridized downstream of the *lef-8* stop codon. The amplified fragment was first cloned into the pCRII vector (Invitrogen) and then subcloned into the T7 expression vector pET15b (Novagen). The resulting plasmid, pET-lef8, was used to express LEF-8 in *Escherichia coli* BL21(DE3)LysE cells. Overexpressed LEF-8 protein was purified by SDS-PAGE and used to generate polyclonal antiserum in mice by using a standard immunization protocol (17).

Samples for immunoblot analysis were boiled for 3 min and electrophoresed on an 8% acrylamide gel. The proteins were electrophoretically transferred to nitrocellulose sheets by using a semidry apparatus. The sheets were reacted with LEF-8 antiserum, and immune complexes were detected by using alkaline phosphatase-conjugated anti-mouse immunoglobulin G.

Primer extension mapping of *lef-9*. Total RNA was isolated from AcNPV-infected cells by the guanidine isothiocyanate-cesium chloride method (10). A *lef-9* specific primer (GTGAGGGTCTAATATGAGG) was radiolabeled at the 5′ end and hybridized with 20 μg of RNA. Annealed primers were extended with avian myeloblastosis virus reverse transcriptase (27). Reaction products were analyzed on 6% polyacrylamide–8 M urea gels. Sequencing ladders were generated by using pPstI-H DNA (14) and the same oligonucleotide primer.

Construction of vBAC-RNApol. The transfer vector pBAC4x-1 (Novagen) contains two copies of the *polh* promoter and two copies of the *p10* promoter with unique restriction sites downstream of each promoter. The four RNA polymerase subunit genes were cloned into this plasmid for overexpression of AcNPV RNA polymerase in infected cells according to standard cloning protocols (27). A *Bam*HI site (GTGCGCAGTAATGGATCCACGATGACGGAC; *Bam*HI site underlined) was inserted upstream of the *lef-8* open reading frame (ORF) in the genomic clone pEcoRI-M by site-directed mutagenesis (11). The resulting 2-kb *Bam*HI-*Eco*RI fragment containing *lef-8* was cloned into pBAC4x-1 (Novagen) under polyhedrin control. The AcNPV genomic clone pHindIII-C was digested with *Nar*I and *Xho*I and incubated with Klenow enzyme and deoxynucleoside triphosphates to fill in 5′ overhangs. A 1.6-kb fragment containing the complete *lef-4* ORF was purified by agarose gel electrophoresis and cloned into the *Sma*I site of pBAC4x-lef8. Insertion of the fragment in the correct orientation was determined by restriction digestions. Site-directed mutagenesis was used to construct a *Bam*HI site (underlined) upstream of the *lef-9* ORF (ACGCGTTCGTGTACGGATCCAAACATGTTT). The resulting plasmid was digested with *Bam*HI and *Bgl*II, and the ends were repaired with Klenow enzyme. A 1.4-kb fragment containing *lef-9* was cloned into the *Stu*I site of pBAC4x-lef8/lef4 under the control of the *polh* promoter. The orientation of the insert was screened by restriction digest analysis. The genomic clone pPstI-F was digested with *Eco*RI and *Pst*I. A 2.3-kb fragment containing p47 was cloned into the *Eco*RI and *Pst*I sites of pVL1393. The p47 ORF was excised from this plasmid by digestion *Bgl*II, followed by repair with Klenow enzyme and digestion with *Bam*HI. The *p47* fragment was then ligated with pBAC-lef8/lef4/lef9 previously prepared by digestion with *Bsu*36I, repair with Klenow enzyme, and digested with *Bgl*II. The resulting plasmid, pBAC-RNApol, was cotransfected with linearized BakPAK6 (Clontech) DNA into Sf9 cells. Recombinant viruses were amplified and used to prepare nuclear extracts by the same protocol as described above except that cells were harvested at 60 h postinfection and the polymin P concentration was increased to 0.2%.

TABLE 1. Purification of baculovirus RNA polymerase from AcNPV-infected cells

Fraction	Protein (mg)	Activity (U)[a]	Sp act (U/mg)	Recovery (%)	Relative purification
Nuclear extract (25 liters)	241.0	ND[b]	ND	ND	ND
Phenyl-Sepharose	47.0	ND	ND	ND	ND
Heparin	13.5	20,526	1,520	100	1
Mono Q	0.42	9,096	21,657	44	14
Superose 6	0.036	4,433	123,138	22	72

[a] Expressed as picomoles of GMP incorporated into RNA in 30 min at 30°C.
[b] ND, not determined (activity was not linear with respect to protein concentration).

RESULTS

Purification of baculovirus RNA polymerase complex. AcNPV RNA polymerase was purified by chromatography of nuclear extracts on phenyl-Sepharose, heparin-agarose, and Mono Q resin, followed by gel filtration through Superose 6 (Table 1). RNA polymerase activity was tested after each purification step by using a baculovirus promoter-specific assay as previously described (30). The transcription complex eluted in a single peak from each column, indicating that all factors required for enzymatic activity and promoter recognition were tightly associated in a single complex. In the crude extract and in the phenyl-Sepharose peak, transcription activity was not linear with respect to amount of protein added, probably because of contaminating nucleases or other proteins that interfere with the assay. Therefore, it was necessary to calculate the purification and yield relative to the heparin peak, and as a result, the values shown in Table 1 underestimate the overall purification. The specific activity of the purified RNA polymerase was approximately 123,000 U/mg of protein.

The distribution of transcription activity on Superose 6 coincided with a peak of protein as measured by absorbance at 280 nm (Fig. 1A). RNA polymerase fractionated with an apparent molecular weight of 560,000. When fractions across the peak were assayed by SDS-PAGE, four polypeptides with apparent molecular weights of 98,000, 55,000, 53,000, and 46,000 were found to increase and decrease concomitant with the peak of enzymatic activity as well as with the peak of protein (Fig. 1B). Additional bands were detected in the molecular weight range of 55,000 to 65,000. These probably represent contaminating keratins since they were observed in most lanes, including lanes with no protein loaded.

The stoichiometry of the four proteins was determined to be equimolar on the basis of Coomassie brilliant blue staining relative to protein markers of known concentration (Fig. 1C), which suggests that an active transcription complex contains two molecules of each subunit. There was no evidence for additional protein bands in the Coomassie blue-stained gel,

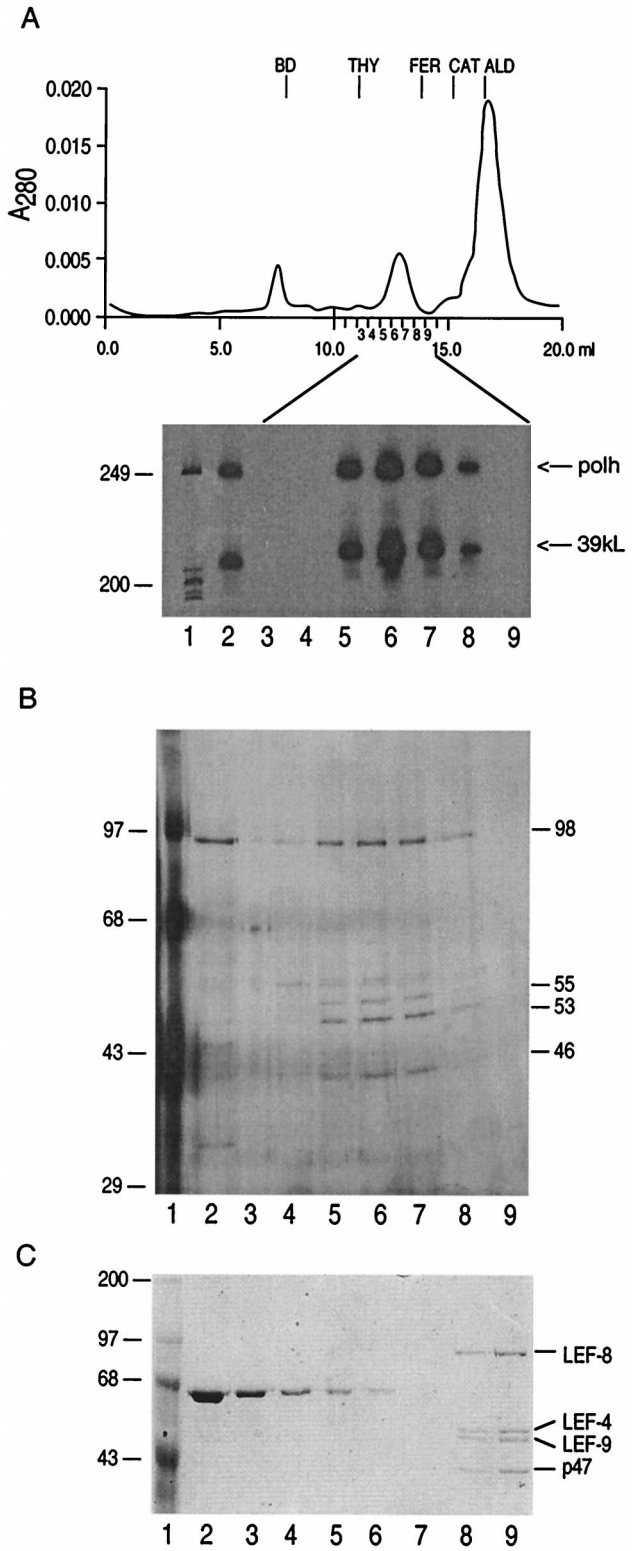

FIG. 1. (A) Gel filtration chromatography of RNA polymerase. RNA polymerase was filtered through Superose 6. Fractions (0.5 ml) were collected from 10 to 15 ml and assayed for transcription activity. The indicated fractions were assayed for in vitro transcription activity as described in Materials and Methods. A transcription assay of the pooled Mono Q peak is shown in lane 2. Lane 1 contains ϕX174-*Hin*fI molecular markers, and the sizes (in kilobases) of relevant fragments are shown on the left. The transcripts corresponding to Polh/CFS and 39kL/CFS are indicated on the right. The positions of elution of blue dextran 2000 (BD), thyroglobulin (THY; 669 kDa), ferritin (FER; 4043 kDa), catalase (CAT; 232 kDa), and aldolase (ALD; 158 kDa) were determined by elution of protein standards under the same conditions. (B) Proteins in the corresponding fractions were separated by electrophoresis on an SDS–8% polyacrylamide gel and visualized by staining with silver. The apparent molecular weights (in kilodaltons) of the bands that correspond with transcription activity are shown on the right. The migration of protein molecular weight markers (in kilodaltons) are indicated on the left. (C) Quantitation of subunits. Purified RNA polymerase (lanes 8 and 9 contain 300 and 150 ng, respectively) was separated by SDS-PAGE and stained with Coomassie brilliant blue. Lanes 2 to 6 contain bovine serum albumin (100, 200, 400, 600, and 800 ng); lane 1 contains molecular markers.

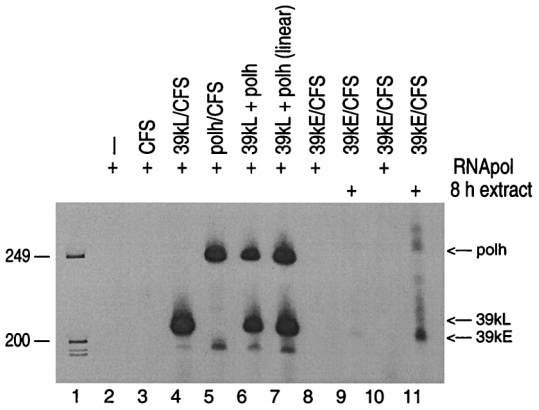

FIG. 2. Template specificity of the very late transcription complex. To analyze transcription of late and very late templates, in vitro transcription assays were performed with 30 ng of purified RNA polymerase (RNApol) and the indicated templates. Plasmids Polh/CFS and 39kL/CFS contain the *polh* promoter linked to a short cytidine-free template as previously described (30). Plasmid 39kE/CFS contains the short cytidine-free template linked to the early *39k* promoter. The transcripts corresponding to Polh, 39kL, and 39kE are indicated on the right. The 39kE/CFS template was transcribed by using purified RNA polymerase (lanes 8 and 10) or with extracts prepared during the early phase of viral infection (lanes 9 and 11) and assayed under conditions optimal for transcription of late templates (lanes 8 and 9) or early templates (lanes 10 and 11). ϕX174 DNA digested with *Hin*fI was added to lane 1, and the sizes (in kilobases) of relevant fragments are indicated on the left.

which suggests that contaminating proteins, if present, were submolar. Analysis of polymerase subunits on a 13% polyacrylamide gel, followed by staining with either silver or Coomassie blue, failed to reveal the presence of smaller subunits (data not shown).

Template specificity of the baculovirus RNA polymerase. In vitro transcription reactions were performed with 30 ng of purified protein and the indicated DNA templates (Fig. 2). No transcripts were detected in the absence of template (lane 2) or in the presence of pCFS (30), the parental cytidine-free cassette (lane 3). When the AcNPV late promoter construct (39kL/CFS) and the very late promoter construct (Polh/CFS) were separately added to transcription reactions, both templates were accurately transcribed (lanes 4 and 5). Both supercoiled and linear templates were transcribed by the RNA polymerase complex with essentially equal efficiencies (lanes 6 and 7).

The purified complex was also not active on 39kE/CFS, a construct that contains the cytidine-free cassette linked to the early promoter for the *39k* gene (16). This early promoter is a typical RNA polymerase II promoter and contains a TATA box and a CAGT initiator element. 39kE/CFS is transcribed by nuclear extracts prepared from AcNPV-infected insect cells harvested at 8 h postinfection (lanes 9 and 11). Transcription of this template was more efficient with conditions optimal for RNA polymerase II templates (6 mM Mg^{2+} and a 15-min preincubation) than with conditions optimal for viral late templates (2 mM Mg^{2+} and no preincubation). Transcripts that

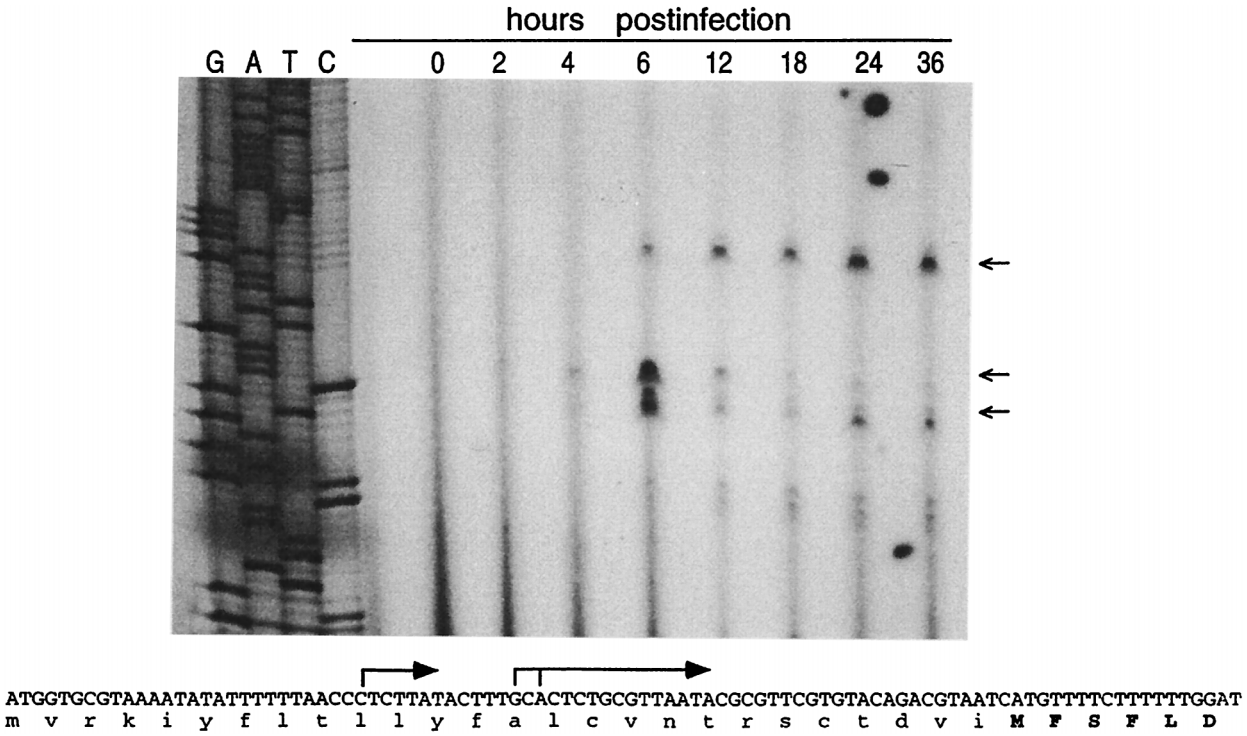

FIG. 3. Primer extension mapping of *lef-9* mRNA. Total cellular RNA was isolated from AcNPV-infected Sf9 cells at the indicated times postinfection. The 5′ end of the transcripts was mapped by primer extension analysis using an oligonucleotide complementary to nt 49300 to 49319 of the AcNPV genome sequence (1). Sequencing ladders were generated by using the same primer. The sequence ladder is antisense relative to the sequence of the *lef-9* promoter shown below. The primer extension products are denoted by arrows on the right and correspond to arrows above the sequence, which indicate the transcription start sites. The amino acid residues identified by N-terminal sequencing are shown in uppercase; the 26-residues predicted to be at the N terminus of LEF-9 are shown in lowercase.

are longer than expected are commonly seen with crude nuclear extracts, presumably due to the presence of contaminating CTP. 39kE/CFS was not transcribed by the viral RNA polymerase under standard conditions for late (lane 8) or for early promoters (lane 10). These data demonstrate that the four-subunit RNA polymerase contains elements necessary for both promoter recognition and enzymatic activity.

Identification of protein subunits. The two smallest RNA polymerase subunits were identified by N-terminal amino acid sequence. Purified RNA polymerase subunits were separated by electrophoresis on SDS–8% polyacrylamide gels, transferred to polyvinylidene difluoride membranes, and submitted for automated Edman degradation. The sequence of the smallest subunit was Met-Phe-Val-Thr-Arg-Leu. The only perfect match for this sequence in the combined SwissProt/GenBank protein databases is the AcNPV protein known as p47. The gene encoding p47 was originally identified as the site of the temperature-sensitive mutant *ts*317 (8). This mutant was defective in the release of infectious virus and expression of polyhedrin at the nonpermissive temperature, although viral DNA synthesis appeared to be normal. This phenotype suggested that the mutation disrupted a function required for transcription of late and very late genes. p47 was also identified by Todd et al. (29) as one of the 18 genes required for transient expression of reporter genes under the control of baculovirus late and very late promoters. The *p47* gene is predicted to encode a protein of 47.5 kDa, which is in good agreement with an apparent molecular mass of 46 kDa calculated from SDS-gels (Fig. 1). Thus, identification of p47 as one of the RNA polymerase subunits is consistent with the known biology of the protein.

The N-terminal sequence of the next-smallest subunit was determined to be Met-Phe-Ser-Phe-Leu-Asp. The only perfect match to this sequence in the protein databases was the AcNPV LEF-9 protein. The *lef-9* gene was originally mapped by Lu and Miller (19) as part of a screen for viral genes required for transient expression of baculovirus late and very late genes. The amino acid sequence that we determined corresponds to residues 27 to 32 of the LEF-9 ORF as originally published (19). However, in that report, transcriptional mapping of the *lef-9* gene was not performed and the ORF was assumed to start at the furthest upstream methionine codon. This discrepancy suggests either that the protein we sequenced was subject to posttranslational processing or the incorrect start site was identified by Lu and Miller (19).

To address the question of which methionine codon was used for initiation of translation of the *lef-9* ORF, we performed primer extension assays. RNA purified from AcNPV-infected cells was hybridized with an oligonucleotide complementary to a sequence within the N terminus of ORF-9. We found that the 5′ ends of *lef-9* transcripts mapped to heterogeneous sites between the upstream methionine codon and the residue identified in our sequence analysis (Fig. 3). Transcription initiating from two start sites proximal to the *lef-9* ORF was detected primarily at the 6-h time point, although low levels of transcripts could be detected at earlier and later times. The proximal transcripts initiated at a CACT motif, which occurs 26 nucleotides downstream of a TATA sequence. This arrangement of promoter motifs is similar to that found in many of the baculovirus early promoters, in which initiation begins at a CAGT motif (4). A single distal transcription start site was mapped to a point 12 nucleotides further upstream. Sequences surrounding this start site do not correspond to consensus early promoters. Transcription from this point was

TABLE 2. MALDI-TOF analysis of LEF-4 tryptic fragments

Mass (Da)	*D* (Da)[a]	LEF-4 peptide sequence consistent with mass	LEF-4 amino acids
839.3	0.29	FVYWPK	77–82
1,057.3	1.1	FEHVYFSK	127–134
1,282.9	−0.06	DTNALVPLLVWR	83–93
1,352.7	0.1	IYITDLLQVFK	306–316
1,640.6[b]	1.6	IVYGDDAFDNASVKK	237–251
1,696.4	−0.05	WMPTTELEYDAVNK	402–415
1,820.3	0.3	EISYSINFSQDLLYK	11–25
1,994.0[b]	0.7	NFCIIQTDDMQFYKTK	268–283
2,887.7	0.5	LLTLLENGDASETLQNSQ VGSDEILAR	151–177
3,142.7	−0.7	ILNSYIVPNYSLAQQYFDL YDENGFR	26–51
3,536.4	1.5	FQQFFDPPLQQSNYMTVS VDGYVVLDTELR	367–396
1,434.2, 1,476.6, 1,536.6, 2,130.9, 2,386.2, 2,707.5, 2,721.0[c]			

[a] Difference between measured mass and calculated mass of LEF-4 tryptic peptides.
[b] Partially cleaved fragments.
[c] Masses not attributed.

detected in at 6 h postinfection, and transcripts persisted through 36 h postinfection.

Our primer extension data strongly argue that the LEF-9 ORF initiates with the sequence Met-Phe-Ser-Phe-Leu-Asp and not at the upstream site previously identified (19). Recalculation of the molecular weight for LEF-9 according to our protein sequence data predicts a polypeptide of 56 kDa, which is closer to the apparent molecular weight of this subunit than the previously published value of 59 kDa.

N-terminal sequence analysis of the 55-kDa polypeptide failed to yield useful sequence information. Therefore, protein fingerprinting was used to identify this subunit. An SDS-polyacrylamide gel slice containing the 55-kDa subunit was digested with trypsin, and then the eluted peptides were analyzed by mass spectroscopy to determine their molecular masses. The masses of all tryptic peptides were entered into the MOWSE database searching program (21), which matches peptide fingerprints of an unknown protein with the predicted fingerprints of all ORFs in the OWL database. The most significant match returned by the database search was for the AcNPV LEF-4 protein. Eleven of the 19 tryptic peptides entered matched those predicted for LEF-4, which confirms the identity of this protein (Table 2). LEF-4 was first mapped by Passarelli and Miller (22) as a factor required for expression of viral late genes, and it was subsequently identified as the site of a temperature-sensitive mutation producing a phenotype similar to the p47 mutation described above (7). The predicted molecular size of LEF-4 is 54 kDa, consistent with an apparent molecular size of 55 kDa calculated from Fig. 1B.

Attempts to sequence the largest subunit were not successful. However, we were able to predict and confirm the identity of the protein by immunochemical analysis. AcNPV encodes relatively few proteins in the molecular weight range observed for the largest subunit. One of these, LEF-8, contains a conserved sequence motif that is common to prokaryotic and eukaryotic RNA polymerases (23). In addition, LEF-8 is required for viral late and very late gene expression in a transient assay, consistent with its proposed role as a component of the virus-specific RNA polymerase. Therefore, we raised antiserum against LEF-8 expressed in bacteria and immunoblotted fractions across the Superose 6 peak. As shown in Fig. 4, the large subunit of purified RNA polymerase was recognized by the LEF-8 antiserum. The intensity of the immunoreactive bands closely correlated with the peak of transcription activity. We conclude, therefore, that the large subunit is encoded by *lef-8*.

Cloning and overexpression of RNA polymerase. To confirm that the RNA polymerase subunits had been correctly identified, we cloned the genes encoding p47, LEF-9, LEF-4, and LEF-8 into the baculovirus transfer vector pBAC4x-1. This plasmid contains two copies of the *polh* promoter and two copies of the *p10* promoter and allows for overexpression of four proteins in a baculovirus expression system. Sf9 cells were infected with the recombinant virus pBAC-RNApol, and nuclear extracts were prepared from cells harvested at 60 h postinfection. In a wild-type virus infection, in vivo synthesis of viral RNAs had ceased by 60 h postinfection, and RNA polymerase activity was undetectable in extracts prepared from wild-type virus-infected cells harvested at this time (data not shown). However, in cells infected with pBAC-RNApol, transcription activity was high in extracts prepared at 60 h postinfection. Overall yields of RNA polymerase per liter of infected cells were 10-fold higher than yields with wild-type virus harvested at 36 h postinfection (Table 3). From 5 liters of cells infected with pBAC-RNApol we were able to purify 70 μg of RNA polymerase, while the yield from wild-type virus-infected cells was only 36 μg from 25 liters of infected cells.

DISCUSSION

Since the discovery of a novel RNA polymerase activity in baculovirus-infected cells (12, 15), the question of whether the

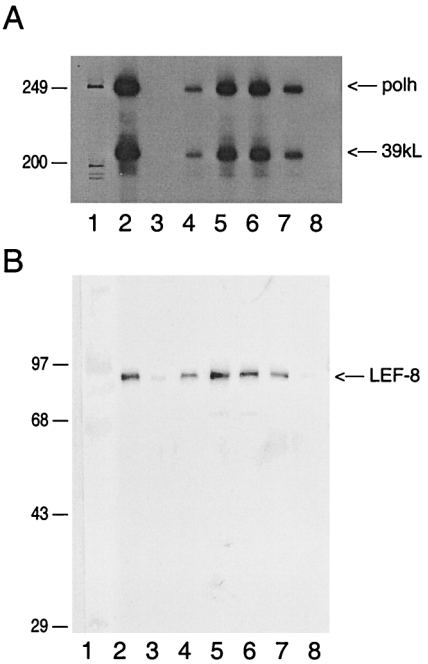

FIG. 4. Recognition of the large subunit with LEF-8 antiserum. (A) The Mono Q peak fraction (lane 2) was filtered through Superose 6 as described in the legend to Fig. 1. Proteins corresponding to the peak of protein at 560,000 were assayed for transcription activity (lanes 3 to 8). Positions of the *polh* and *39k*L transcripts are indicated on the right. ϕX174 molecular markers are shown in lane 1. (B) Superose 6 fractions containing transcription activity were electrophoresed on SDS–8% polyacrylamide gels, transferred to nitrocellulose membranes, and probed with LEF-8 antiserum. Lane 1, prestained molecular weight markers, with sizes (in kilodaltons) indicated on the left. The position of the immunoreactive protein is shown on the right.

TABLE 3. Purification of baculovirus RNA polymerase from pBAC-RNApol-infected cells

Fraction	Protein (mg)	Activity (U)[a]	Sp act (U/mg)	Recovery (%)	Relative purification
Nuclear extract (5 liters)	50	ND[b]	ND	ND	
Phenyl-Sepharose	14.8	48,263	3,261	100	1
Heparin	2.4	44,754	18,648	93	5.7
Mono Q	0.26	15,019	57,765	31	17.7
Superose 6	0.07	10,175	145,357	21	44.6

[a] Expressed as picomoles of GMP incorporated in 10 min at 30°C.
[b] ND, not determined.

α-amanitin-resistant polymerase is a virus-encoded enzyme or a virus-modified host RNA polymerase has intrigued baculovirologists. This report identifying the protein subunits of the AcNPV polymerase confirms that the polymerase responsible for transcription of late and very late genes is encoded by the virus.

The baculovirus RNA polymerase was shown to consist of only four subunits, p47, LEF-4, LEF-8, and LEF-9. The specific activity of the polymerase purified from wild-type virus-infected cells indicated that the enzyme could incorporate approximately 123 nmol of GMP per mg of enzyme. Assuming a monomer molecular weight of 258,000, this implies that 1 nmol of enzyme can incorporate 32 nmol of GMP. The cytidine-free template directs incorporation of 90 GMP residues, indicating that 35% of the RNA polymerase complexes are active in transcription. If the viral RNA polymerase contains two molecules of each subunit as suggested by the gel filtration data, this would suggest that 70% of the molecules are active. This high molar activity of the purified enzyme suggests that the transcription activity is due solely to the four proteins detected. While we cannot rule out the possibility that our purified polymerase preparations contain submolar amounts of other viral or host proteins, these calculations suggest that it is unlikely that proteins present at submolar levels could significantly contribute to activity. However, definitive proof that these four proteins are sufficient for RNA polymerase activity will require reconstitution of activity from purified subunits.

Overexpression of the four subunits in baculovirus-infected cells was shown to increase the yields of purified polymerase by a factor of 10. This confirms our identification of the four subunits and suggests that one or more of these proteins are limiting for the assembly of RNA polymerase. Furthermore, the specific activity of RNA polymerase purified from the overexpression system was equivalent to that from wild-type-infected cells. This finding supports our conclusion that minor contaminating proteins do not contribute significantly to activity, as it is unlikely that these minor proteins would contaminate the overexpressed preparation at levels 10-fold higher than with enzyme purified from wild-type virus-infected cells.

All four of the polymerase subunits have previously been shown to be required for transient expression of viral late and very late genes (29), and genetic evidence suggested that *p47* and *lef-4* encode late transcription factors (7). LEF-8 and LEF-9 were predicted to encode RNA polymerase subunits, based on the presence of conserved motifs (19, 23). This LEF-8 motif is conserved among β and β′ subunits of RNA polymerases from a number of sources and is predicted to form part of the catalytic site of the enzyme. Our data showing that LEF-8 is the largest subunit in the purified RNA polymerase complex support this hypothesis, although we have not yet determined whether LEF-8 is the catalytic subunit.

The baculovirus RNA polymerase was active only on templates containing a viral late promoter. Cytidine-free templates lacking a promoter or linked to an early promoter were not transcribed, suggesting that the conserved TAAG motif was essential for transcription initiation. Late and very late promoters were transcribed with equal efficiencies, suggesting that additional factors are needed for correct temporal expression of the two classes of late genes. In vivo, baculovirus late and very late genes are expressed during two distinct temporal phases. The late genes are transcribed first, and only the sequences immediately surrounding the TAAG motif are believed to be important in promoter selection. The very late genes are expressed after assembly of virions, and in addition to the essential TAAG motif, an A+T-rich region between the start of transcription and the start of translation is important for high-level expression of the very late genes (25). Thus, the polymerase complex that we purified is probably lacking the factor(s) responsible for the burst of very late transcription, and possibly it is also missing the factor(s) that represses polyhedrin transcription during the late phase. Previously we showed that phosphocellulose chromatography of extracts prepared from infected cells could separate factors that altered the template specificity of the viral RNA polymerase (30). While phosphocellulose chromatography may be useful in the future for purification of additional transcription factors, we found it necessary to omit this step from our current protocol. The RNA polymerase complex aggregates at low salt, and significant amounts of polymerase were lost during dialysis prior to loading onto phosphocellulose.

Previous work has identified 18 viral proteins that are essential for expression of late and very late genes (29). Eight of these are required for early viral gene expression and DNA replication which precede late gene expression. Therefore, a model for late gene expression should include roles for the remaining 10 proteins. Here we show that 4 of these 10 are components of the viral RNA polymerase. Functions of the additional proteins probably include factors that discriminate between late and very late genes, as mentioned above. In addition, proteins are needed for termination of transcription, a function which would not be identified in our assay because transcription stalls when a cytidine is needed. Furthermore, posttranslational processing functions (formation and methylation of 5′ caps, cleavage, and polyadenylation, for example) are probably virus encoded as well. It is also possible some of these viral proteins are involved in assembly of the four polymerase subunits. Since our overexpression was done in a baculovirus system, all other viral proteins should be present in the infected cells and could act steps that precede initiation of transcription.

Most eukaryotic DNA viruses rely on the host RNA polymerase II for transcription of viral genes. Poxviruses and African swine fever virus are the only other examples of eukaryotic viruses that encode DNA-dependent RNA polymerases, and their need for a virus-encoded enzyme is usually attributed to the fact that they replicate in the cytoplasm and thus do not have access to the nuclear enzymes. The poxvirus enzymes resemble the three cellular RNA polymerases with respect to the number and size of subunits (5). All of the cellular enzymes contain two large subunits which share regions of amino acid sequence homology and 4 to 12 small subunits, some of which are common to the different cellular enzymes and homologous to the poxvirus enzymes. The AcNPV RNA polymerase with only four subunits is the simplest DNA-directed RNA polymerase from a eukaryotic source described thus far. The baculovirus RNA polymerase contains only one large subunit, and with the exception of two short motifs it shows little homology with other RNA polymerases. This lack of similarity with re-

spect to sequence and structure raises interesting questions regarding the evolutionary origin of the baculovirus RNA polymerase.

Our data suggest that baculoviruses are more similar to T7 and related bacteriophages than to the host or other eukaryotic viruses. The replication strategies employed by the two viruses are strikingly similar, although one infects insect cells and the other infects bacteria. In both cases, the early viral promoters resemble their host promoters and are transcribed by the corresponding host RNA polymerases. However, the structures of the viral late gene promoters are dramatically different from those of their early promoters and they do not contain motifs recognized by the host polymerases. In fact, the structures of the baculovirus and T7 late promoters are more similar to each other than either is to the structures of the promoters of their hosts. Both promoters consist of a short conserved sequence that serves as both a promoter and an initiator element (25, 26). Therefore, it is not surprising to find that baculoviruses, like T7 bacteriophage (9), use a virus-encoded RNA polymerase to transcribe viral late genes. The T7 RNA polymerase is a single polypeptide which contains both promoter recognition and enzymatic activities (2). Although simple in structure, the viral polymerases transcribe their cognate genes efficiently and with high specificity.

ACKNOWLEDGMENTS

This research was supported by grant MCB 95-06233 from the National Science Foundation. Mass spectrometry and automated sequence analysis were performed at California Institute of Technology with equipment purchased under NIH grant RR11292.

REFERENCES

1. **Ayres, M. D., S. C. Howard, J. Kuzio, M. Lopez-Ferber, and R. D. Possee.** 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. Virology **202:**586–605.
2. **Bailey, J. N., J. F. Klement, and W. T. McAllister.** 1983. Relationship between promoter structure and template specificities exhibited by the bacteriophage T3 and T7 RNA polymerases. Proc. Natl. Acad. Sci. USA **80:**2814–2818.
3. **Beniya, H., C. J. Funk, G. F. Rohrmann, and R. F. Weaver.** 1996. Purification of a virus-induced RNA polymerase from *Autographa californica* nuclear polyhedrosis virus-infected Spodoptera frugiperda cells that accurately initiates late and very late transcription in vitro. Virology **216:**12–19.
4. **Blissard, G. W., and G. F. Rohrmann.** 1990. Baculovirus diversity and molecular biology. Annu. Rev. Entomol. **35:**127–155.
5. **Broyles, S. S., and B. Moss.** 1987. Homology between RNA polymerases of poxviruses, prokaryotes, and eukaryotes: nucleotide sequence and transcriptional analysis of vaccinia virus genes encoding 147-kDa and 22-kDa subunits. Proc. Natl. Acad. Sci. USA **83:**3141–3145.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
7. **Carstens, E. B., H. Chan, H. Yu, G. V. Williams, and R. Casselman.** 1994. Genetic analysis of temperature-sensitive mutations in baculovirus late expression factors. Virology **204:**323–337.
8. **Carstens, E. B., A. L. Lu, and H. L. B. Chan.** 1993. Sequence, transcriptional mapping, and overexpression of p47, a baculovirus gene regulating late gene expression. J. Virol. **67:**2513–2520.
9. **Chamberlin, M., and J. Ring.** 1973. Characterization of T7-specific ribonucleic acid polymerase. 1. General properties of the enzymatic reaction and the template specificity of the enzyme. J. Biol. Chem. **248:**2235–2244.
10. **Chirgwin, J. M., A. E. Przybyla, R. J. McDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294–5297.
11. **Deng, W. P., and J. A. Nickoloff.** 1992. Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. Anal. Biochem. **200:**81–84.
12. **Fuchs, L. Y., M. S. Woods, and R. F. Weaver.** 1983. Viral transcription during AcNPV infection: a novel RNA polymerase induced in infected *Spodoptera fugiperda* cells. J. Virol. **48:**641–646.
13. **Glocker, B., R. R. Hoopes, and G. F. Rohrmann.** 1993. In vitro transcription from baculovirus late gene promoter: accurate mRNA initiation by nuclear extracts prepared from infected *Spodoptera frugiperda* cells. J. Virol. **67:**3771–3776.
14. **Gong, M., and L. A. Guarino.** 1994. Expression of the 39k promoter of *Autographa californica* nuclear polyhedrosis virus is increased by the apoptotic suppressor P35. Virology **204:**38–44.
15. **Grula, M. A., P. L. Buller, and R. F. Weaver.** 1981. α-Amanitin-resistant viral RNA synthesis in nuclei isolated from nuclear polyhedrosis virus-infected *Heliothis zea* larvae and *Spodoptera frugiperda* cells. J. Virol. **38:**916–921.
16. **Guarino, L. A., and M. Smith.** 1992. Regulation of delayed-early gene transcription of dual TATA boxes. J. Virol. **66:**3722–3739.
17. **Harlow, E., and D. Lane.** 1988. Antibodies: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Hoopes, R. R., Jr., and G. F. Rohrmann.** 1991. In vitro transcription of baculovirus immediate early genes: accurate initiation by nuclear extracts from both insect and human cells. Proc. Natl. Acad. Sci. USA **88:**4513–4517.
19. **Lu, A., and L. K. Miller.** 1994. Identification of three late expression factor genes within the 33.8- to 43.4-map-unit region of *Autographa californica* nuclear polyhedrosis virus. J. Virol. **68:**6710–6718.
20. **Mans, R. M., and D. Knebel-Morsdorf.** 1998. In vitro transcription of pe38/polyhedrin hybrid promoters reveals sequences essential for recognition by the baculovirus-induced RNA polymerase and for the strength of very late viral promoters. J. Virol. **72:**2991–2998.
21. **Pappin, D. J., P. Hojrup, and J. J. Bleasby.** 1993. Rapid identification of proteins by peptide-mass fingerprinting. Curr. Biol. **3:**327–332.
22. **Passarelli, A. L., and L. K. Miller.** 1993. Identification of genes encoding late expression factors located between 56.0 and 65.4 map units of the Autographa californica nuclear polyhedrosis virus genome. Virology **197:**704–714.
23. **Passarelli, A. L., J. W. Todd, and L. K. Miller.** 1994. A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases. J. Virol. **68:**4673–4678.
24. **Pullen, S. S., and P. D. Friesen.** 1995. Early transcription of the *ie1* transregulator gene of *Autographa californica* nuclear polyhedrosis virus is regulated by DNA sequences within its 5′ noncoding leader region. J. Virol. **69:**156–165.
25. **Rankin, C. B., B. G. Ooi, and L. K. Miller.** 1988. Eight base pairs encompassing the transcriptional start point are the major determinant for baculovirus polyhedrin expression. Gene **70:**39–49.
26. **Rohrmann, G. F.** 1986. Polyhedrin structure. J. Gen. Virol. **67:**1499–1508.
27. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
28. **Summers, M. D., and G. E. Smith.** 1987. A manual of methods for baculovirus vectors and insect cell culture procedures. Tex. Agric. Exp. Stn. Bull. 1555.
29. **Todd, J. W., A. L. Passarelli, and L. K. Miller.** 1995. Eighteen baculovirus genes, including *lef-11*, *p35*, *39k*, and *p47*, support late gene expression. J. Virol. **69:**968–974.
30. **Xu, B., S. Yoo, and L. A. Guarino.** 1995. Differential transcription of baculovirus late and very late promoters: fractionation of nuclear extracts by phosphocellulose chromatography. J. Virol. **69:**2912–2917.
31. **Yang, C. L., D. A. Stetler, and R. F. Weaver.** 1991. Structural comparison of the *Autographa californica* nuclear polyhedrosis virus-induced RNA polymerases from the host, *Spodoptera frugiperda*. Virus Res. **20:**251–264.
32. **Yoo, S., and L. A. Guarino.** 1994. The *Autographa californica* nuclear polyhedrosis virus *ie2* gene encodes a transcriptional regulator. Virology **202:**746–753.